Guilt aversion and other motivations: Eve versus Adam*

Giovanni Di Bartolomeo,[a] Martin Dufwenberg,[b] Stefano Papa,[c] and Laura Razzolini[d]

January 31, 2023

Abstract. We explore gender differences in how people are motivated. We focus on guilt aversion, a surprisingly relatively unexplored issue. Our experiment supports the idea that men are more guilt averse than women. We provide a potential intuition for our findings based on the pregnancy-related biological asymmetry between genders. Finally, our paper explores other rationales to explain observed gender behaviors, like moral commitment.

***JEL* codes:** A13, C91, D03, D64.
Keywords: gender, guilt aversion, promises, evolutionary psychology

1. Introduction

Evolutionary psychologists Tooby and Cosmides (2008, p. 177) argue that "each emotion evolved to deal with a particular, evolutionarily recurrent situation." This statement may suggest that emotions affect men and women differently if the games they play cast the sexes in asymmetric roles. A case in hand may concern *guilt*. Men may have more substantial evolutionary motivation to be known to be prone to guilt for child-bearing reasons. A pregnant mother spends nine months gestation when the father could conceivably take off and produce offspring with other women. If women anticipated such opportunistic behavior, they might not agree to conceive in the first place.

---

* We are grateful to Giuseppe Attanasi and Leda Cosmides for helpful discussions.

[a] Department of Economics and Law, Sapienza University of Rome; CIMEO; Department of Economics, University of Antwerp.
[b] Department of Economics, University of Arizona; Department of Economics, University of Gothenburg; CESifo.
[c] Department of Economics and Finance, University of Rome Tor Vergata; CIMEO.
[d] Department of Economics, Finance and Legal Studies, University of Alabama.

If a prospective father is known to be prone to feelings of guilt, that may prevent him from leaving the mother and, therefore, help secure her trust.

Behavioral and experimental economists have defined and tested guilt aversion (GA) in games and shown that GA can help explain trustworthy behavior in the lab.[1] This literature, however, has not extensively addressed the issue of gender differences in GA. Motivated by the trust-during-pregnancy example, we experimentally explore in the laboratory whether GA affects men differently than women. The directional nature of such an effect is not self-evident, however. While the nine-month gestation story given above seems important, other stories may be provided in support of female, rather than male, guilt-proneness.

Our study contributes to the extensive literature in experimental economics that explores gender differences in how individuals are motivated.[2] We use a simple dictator game with pre-play communication to disentangle the non-monetary (potentially different) motivations that drive women's and men's actions. We distinguish motivations driven by belief-dependent GA, as well as motivations that are independent of beliefs. We refer to the latter as moral commitment, MC. Testing for differences in motivations is a complex task for two reasons: i) GA- and MC-driven motivations are not mutually exclusive; ii) different motivations imply different causations, but they are compatible with the same correlations between actions and expectations.

We propose a set of tests based on exogenous variations and difference-in-difference comparisons to face the above-described challenging task. Here, we describe our approach to grasp intuition; more details will be introduced later in the following sections.

In our experiment, pairs communicate with each other and exchange promises about their potential altruistic behavior in a dictator game. After communicating, the pair plays the dictator game, with one individual chosen to be the dictator and the other to be the recipient. However, with a given known probability, some pairs are switched, so that a dictator can end up playing with a recipient different from the one with whom he had previously communicated. Furthermore, the

---

[1] Battigalli and Dufwenberg (2007) develop a general theory. Dufwenberg and Gneezy (2000) and Charness and Dufwenberg (2006) ran early experiments. See, among others, Khalmetski et al. (2015), Bellemare *et al*. (2017), Ederer and Stremitzer (2017), Attanasi *et al*. (2019), Dhami *et al*. (2019), Di Bartolomeo *et al*. (2019a, 2019b, 2022) and Cartwright (2019) who survey the subsequent experimental literature. See Battigalli and Dufwenberg (2022, especially Sections 3.1 and 7) for a broad related discussion.

[2] For pioneering efforts, see Bolton and Katok's (1995) and Eckel and Grossman's (1998) studies of generosity. Croson and Gneezy (2009) surveyed many papers that explored risk preferences, social preferences, or attitudes to competition. Other examples include Dreber and Johannesson (2008) and Childs (2012), who report results regarding attitudes to lying.

switch is revealed only to the dictators, so the recipients' expectations, their first-order beliefs (FOBs), and the dictators' second-order beliefs (SOBs) about the recipients' beliefs are independent of the switch.

Promises fuel expectations independently of the driving motivations. Therefore, on average, recipients who experienced promises would have higher expectations when the switching probability is low than those who experienced promises in a high switching probability scenario. This way, we obtain exogenous variation in beliefs with high or low SOBs, as dictators know that recipients do not observe whether the switch occurred. Different implications for GA can, then, be searched in the data. Similarly, we can glean variation in intrinsic motivations by comparing, everything else equal, the behavior of a dictator asked to keep his/her own promise to the behavior of a dictator asked to keep the promise made by someone else. Here, intrinsic motivations, such as MC, imply that we are more likely to observe promise-keeping by the former dictators than the latter. Once we separately explore MC and GA by gender, we can test how each affects behavior across genders in a difference-in-difference approach. In the case of MC, we focus on its marginal effect, i.e., the difference in the behavior of promisors asked to keep their own promises to that of promisors asked to keep a promise made by someone else. Then, we compare the marginal effects for men and women, respectively. Similarly, the marginal effect of GA is measured by the change in behavior due to a change in SOBs, and then, we test by gender the difference in marginal effects.

Our paper is related to the rich literature on GA (see footnote 1). The non-belief-dependent motivations relevant to our point of view (MC) are discussed and surveyed, among others, by Ellingsen and Johannesson (2004) and Vanberg (2008). Moreover, the following three papers, which are closely related to our work, should be discussed: Vanberg (2008), Di Bartolomeo *et al.* (2019b), and Kleinknecht (2019). We share with these papers our experimental setup, a random mini-dictator game with pre-play communication.

Vanberg (2008), in a pioneering study, introduced the exogenous variation technique previously described to test the effects of MC. He assumed a fixed 0.5 switching probability and compared the behavior of dictators who were asked to keep their own promises to the behavior of dictators who were asked to keep promises made by others. He found evidence for MC. By using different switching probabilities (high and low), Di Bartolomeo *et al.* (2019b) extend Vanberg (2008) to obtain a double exogenous variation (MC and SOBs) to test both GA and MC-driven motivations. They found (some) support for both. Neither Vanberg (2008) nor Di Bartolomeo *et* *al*. (2019b) consider the possibility of gender differences; however, Kleinknecht (2019) replicated Vanberg's (2008) experiment and controlled for the gender of the participants. She did not find evidence for MC for men or women.

While Vanberg's (2008) primary focus was his MC theory of promise-keeping, his approach also permits testing for GA. He explored how the behavior of switched dictators who sent a promise changed depending on whether or not their new recipient had previously received a promise from someone else. He thereby tested for GA because dictators who play with recipients who received a promise hold higher SOBs than those who did not. Vanberg did not find support for GA, while Kleinknecht (2019) did. However, Vanberg's and Kleinknecht's tests have the following questionable feature: They compare the behavior of dictators whose switched recipients received a promise from someone else to the behavior of those of dictators whose switched recipients did not receive a promise. This comparison involves subjects who differ in two ways. First, dictators who play with recipients who received a promise have higher SOBs than those who did not. Second, the comparison involves dictators who play with recipients who received a promise and recipients who did not. A comparison along two dimensions (rather than one) is undesirable from an experimental methodology point of view. The design by Di Bartolomeo *et al.* (2019b) avoids this problem since their comparison involves dictators who play with recipients who differ only in their expectations, not in whether they received a promise or not.

The current paper innovates the above literature in two dimensions. First, we test for gender differences using the Di Bartolomeo *et al*. (2019b) framework. In this respect, we extend Kleinknecht (2019) by considering a double exogenous variation in SOBs and promises. As a result, we cleanly isolate and test for both MC- and GA-driven motivations within and between genders. Our approach allows us to overcome the questionable features of Vanberg's and Kleinknecht's GA tests discussed above. Second, we introduce a methodological innovation by considering a difference-in-difference approach to explore gender sensitivity to GA (or MC.) In this respect, we compare the pro-social incremental effects of GA (or MC) observed in men and women. An observed difference would support the idea of differences in gender sensitivity to GA (or MC.)

Sections 2 and 3 present our experimental design, hypotheses, and procedures. Section 4 shows our main results. Section 5 offers a concluding discussion.

2. The experiment and our hypotheses

2.1 The game

We use a two-player mini-dictator game augmented with a) bilateral pre-play communication, b) roles' randomness and asymmetric information, and c) a partner-switching mechanism. It works in five steps:

1) Participants are matched in pairs. Each pair can communicate by sending messages. The subjects know they will play a mini-dictator game but do not know who will act in which role. Subjects can make promises about their behavior if they are chosen to be dictators.
2) After the communication phase, each player is assigned a role, either dictator or recipient.
3) After the communication phase and before making the allocation decision, a given proportion of dictators have their partners switched. Both players know the proportion of switched pairs, but only dictators are told whether their partner has been switched or not. The recipients are not informed whether a switch occurred.
4) After the switch, dictators can read the messages (sent by another dictator) received by their new partners.
5) Each pair plays the game (form) shown in Figure 1. Dictators choose *Roll* or *Don't Roll*. The payoffs are in Euro, and the recipient's payoff is listed on top. The information set "0···0" indicates that the recipient is not informed how the payoff of "0" arose.

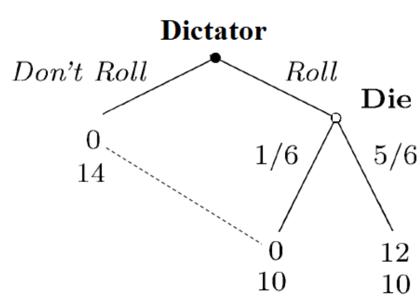

Figure 1 – Game payoffs

2.2 Hypotheses

Our exogenous-variation-based tests involve a subgroup of dictators. Specifically, we focus on dictators who made and "read" promises and refer to this group as **Γ-dictators**. Non-switched Γ-dictators are those who must decide whether to keep their own promises, while switched Γ-dictators are promisors who are re-matched with a recipient who received a promise from another dictator. As recipients are not informed whether a switch occurred, FOBs and SOBs are independent of the switch. However, these beliefs can (and we show that they do) depend on the switching probability: a low (high) switching probability is associated with high (low) beliefs. Unless we explicitly say otherwise, we always refer to Γ-dictators.

2.2.1 SOBs and gender

We begin by introducing our hypotheses about SOBs. We denote by $S_i(j,z,k)$ the average SOBs of a dictator who is $i \in \{M[ale], F[emale]\}$, in a game with $j \in \{L[ow], H[igh]\}$ switch probability, paired with a recipient who received a $z \in \{P[romise], E[mpty\ communication]\}$ message during the pre-play communication (where by empty communication we mean no promise), and who has been $k \in \{S[withed], N[ot\ swithed]\}$.

We start with two preliminary hypotheses crucial to all that follow, since the subsequent hypotheses are conditional on their relevance (Spoiler alert: They will be supported!). First, we explore whether there is any difference in the SOBs between women and men for any given switching probability (H1) and whether our exogenous variations in SOBs work well for both genders (H2). Formally:

H1 (no difference in expectations between genders): Subjects' expectations are independent of gender: $S_M(j,z,k) = S_F(j,z,k)$ for $j \in \{L,H\}$, $z \in \{P,E\}$ and $k \in \{S,N\}$.

H2 (exogenous variations in beliefs within genders): Subjects' expectations are higher when the switching probability is low, for both men and women: $S_i(L,z,k) > S_i(H,z,k)$ for $i \in \{M,F\}$, $z \in \{P,E\}$ and $k \in \{S,N\}$.

H2 is expected to hold on the basis of the extensive literature that provides evidence about a correlation between promise-keeping behavior and SOBs (for related results and discussion, see Vanberg, 2008; Di Bartolomeo *et al.*, 2019b).

2.2.2 Guilt aversion and gender

We focus on switched subjects to explore GA. In this way, we can abstract from any MC motivations for subjects' behaviors and concentrate on GA only. To formalize our hypotheses, we define by $R_i(j, z, k)$ the average *Roll* rate of $i \in \{M[ale], F[emale]\}$ Γ-dictators, who played a game with a $j \in \{L[ow], H[igh]\}$ switch probability and were paired with a recipient who received a $z \in \{P[romise], E[mpty\ communication]\}$ message and who has been $k \in \{S[withed], N[ot\ swithed]\}$. We test two main hypotheses. The f first one is a clear implication of GA. The logic is that, assuming that H2 holds, since a Γ-dictator's SOB is higher when the switching probability is low, if said Γ-dictator is guilt averse, he is more likely to *Roll*. Formally:[3]

H3 (implication of GA within genders): Switched Γ-dictators are more likely to *Roll* when the switching probability is low: $R_i(L, P, S) > R_i(H, P, S)$ for $i \in \{M, F\}$.

We shall test H3 for each gender.

We furthermore explore differences in GA sensitivity between genders. We compare the potential increase in the average *Roll* rate for men associated with a rise in SOBs to the equivalent potential increases for women. Formally, assuming H2 holds, we use difference-in-difference in GA sensitivity between men and women as follows:

H4 (different guilt sensitivities between genders): Male switched Γ-dictators have a higher *Roll* rate than female switched Γ-dictators, when SOBs are higher (low switching probability): $R_M(L, P, S) - R_M(H, P, S) > R_F(L, P, S) - R_F(H, P, S)$.

2.2.3 Moral commitment and gender

To test MC-driven motivations, we need to reintroduce in the analysis non-switched dictators. As in Vanberg (2008) and Kleinknecht (2019), we also compare the behavior of dictators asked to keep their own promises to the behavior of promisor dictators asked to keep a promise made by someone else. Clearly, only the former dictators could be driven by MC motivations. Hence,

[3] It is worth noting that Vanberg's GA test, also used by Kleinknecht (2019), would be $R_i(j, P, S) > R_i(j, E, S)$ for $i \in \{M, F\}$ and $j \in \{L, H\}$. Here SOBs are expected to be different (and they are), but the messages z are also different. In any case, we also perform these tests and report them in Appendix B.

observing a difference in average behavior across the two groups of dictators would provide evidence in favor of MC.[4]

Again, we test two hypotheses. We first explore whether subjects are more likely to keep their own word than promises made by others (H5). As in Kleinknecht (2019), we perform Vanberg's test of MC within gender. However, we do it for different levels of SOBs. This leads to four tests to implement Vanberg (2008)'s test combining gender (women/men) and SOBs (high/low) differences. Assuming that H1 holds, the tests are as follows:

H5 (implication of MC within gender): For a given switching probability, not-switched Γ-dictators are more likely to *Roll* than switched Γ-dictators who read a promise. That is $R_i(j,P,N) > R_i(j,P,S)$ for $i \in \{M,F\}$ and $j \in \{L,H\}$. Combining i and j, we obtain four MC tests.

The next hypothesis is a between-gender test to explore differences in MC sensitivity between genders when SOBs are either high or low. We compare the increase in the average *Roll* rate for men, presumably driven by MC, to the analogous increase for women. In other words, we perform a difference-in-difference comparison in the MC sensitivity between men and women.

We run the following tests, for for each switching probability:

H6 (different moral sensitivities between genders): For a given switching probability, male and female non-switched and switched Γ-dictators who read a promise are equally likely to *Roll*. That is $R_F(j,P,N) - R_F(j,P,S) = R_M(j,P,N) - R_M(j,P,S)$ for $j \in \{L,H\}$.

3. Procedures

The experiment was conducted at the CIMEO Experimental Economics Lab of Sapienza University of Rome. The experiment involved 384 undergraduate student subjects (12 sessions of 8 rounds, with 32 subjects in each session) recruited using an online system. Upon arrival, subjects

[4] A fair comparison needs to consider only switched dictators who also made a promise before the switch, as promisor and non-promisors could have different attitude towards *Rolling*. This is the reason why we focus on Γ-dictators.

were randomly assigned to 32 isolated computer terminals. Three assistants handed out instructions (see the Supplementary Material online) and checked that participants correctly followed the procedures. Before playing any game, subjects completed a short questionnaire testing their comprehension.

Each session consisted of eight rounds, with perfect stranger matching. Payoffs like in Figure 1 were computed in tokens (where one token = 0.50 euro). At the end of each session, one round was randomly chosen for payment. FOBs and SOBs were elicited by asking subjects to guess their counterparts' actions and guesses. Incentives were provided for all rounds except the one chosen for payment, implying that subjects had no incentive to hedge against bad outcomes and, thus, misreport their beliefs.[5] All subjects received a fixed show-up fee of 2.50 tokens.

In each round, the following five stages were implemented:

1. **Communication.** Subjects were randomly matched to form 16 chatting pairs, with a random determination of who would start the chat. As in Vanberg's design, each chat consisted of four one-way messages in sequence. Each message could be of at most 90 characters and was cataloged as involving a promise or not (see below).
2. **Role assignment and revelation of the switching probability.** After the communication stage, roles were randomly assigned in each pair, and subjects were informed of that. Then, depending on the treatment, the switching probability was announced as either 25% (low) or 75% (high).
3. **Belief elicitation**. This stage has two parts:
   a. FOBs: each recipient was asked to guess his/her expected payoff.
   b. SOBs: dictators were asked to guess the payoff expected by the recipient with whom he/she had communicated in the chatting pair.
4. **Switching.** Depending on the treatment, 25% or 75% of recipients were switched. Only dictators were informed whether a switch occurred. Dictators with switched recipients were allowed to read the prior conversation of their new recipient.

[5] Our elicitation procedure is described in detail in Appendix. For the sake of comparison with Di Bartolomeo *et al.* (2019b), we elicited beliefs before the switch occurred and before dictators made their choice.

5. **Dictators' action.** All dictators chose *Roll* or *Don't Roll*. All subjects were informed of their payoff for the round. Recipients were not informed whether they had been switched, nor could they infer the dictator's choice when their payoffs were zero.[6]

Messages were classified according to Vanberg's protocol. Following Vanberg (2008), we refer to each chat sent by a subject in a round as a "message." We had 3,072 messages (32 subjects in 12 sessions of 8 rounds.) We asked two research assistants to code messages according to whether they conveyed a promise or statement of intent indicating that the subject would *Roll*. Ex-ante, we decided to randomly use the code of only one of two assistants. Assistants were unaware of this choice.

4. Main results

Our sample contains 2,240 promises out of 3,072 messages. The promise rates of men and women (71% vs. 75%) are not statistically different (Z=0.47, p=0.638).[7] All statistics are obtained using the Wilcoxon signed rank test, which compares averages at the session level. Our data are independent at the session level, but not at the individual level.[8] We test the hypotheses described in Section 2.2. We first test for gender differences in SOBs and if our exogenous variation works. We presume there is no difference in SOBs between women and men for any given switching probability (H1) and that the exogenous variation works for both genders (H2).

Table 1 reports the SOBs of Γ-dictators. The switching probabilities are reported by row, and the dictators' gender is listed by column. For instance, the value in the first cell of the table means that 73% of women-dictators who did a promise believe that the recipient thinks that the dictator will roll after making a promise when the switching probability is low, formally, average $S_F(L,P,S) = 0.73$. Standard deviations and number of observations are indicated in parentheses.

We test H1 by comparing the SOBs in each row. We do not find any statistically significant difference between men's and women's SOBs, when the chance of being rematched is low, 76%

[6] Recipients could obtain a zero payoff in two cases: (i) their dictator had chosen *Don't Roll*; (ii) their dictator had chosen *Roll,* and the outcome of the die-roll was "1."

[7] Across all subjects, the frequency of promises does not statistically differ across high/low switching probability treatments (i.e., 76% vs. 73%: Z=1.16, p=0.247).

[8] In comparing men versus women, we use two-tailed tests, as we do not have preconceived directional hypotheses. Instead, when we test the rationale of promise-keeping, we use one-tailed tests since we have directional hypotheses.

vs. 73% (Z=1.26, p=0.209), or high, 62% vs. 66% (Z=-0.16, p=0.875). Our results support the non-existence of a gender difference in SOBs (H1).

The data in Table 1 are also consistent with exogenous variations in expectations (H2). Promisors' SOBs reported in column (a) are high (low) when the chance of being rematched is low (high). The result holds for each gender: for women (column (a)): 0.73 vs. 0.66 (Z=2.04, p=0.04); and for men (column (b)) 0.76 vs. 0.62: (Z=3.06, p=0.002).

Table 1 – SOBs of Γ-dictators

TREATMENT	SWITCHING PROBABILITY	**Women** SWITCH (a)	**Men** SWITCH (b)
(i)	25% (low)	**0.73** (0.34/266)	**0.76** (0.27/251)
(ii)	75% (high)	**0.66** (0.35/229)	**0.62** (0.35/234)

We focus next on GA (H3-H4). Table 2 reports the *Roll* rates of switched Γ-dictators when the probability of switching is low or high. As we consider only switched pairs, dictators' motivations cannot be related to MC. Moreover, their SOBs are likely to be high when the switching probability is low and vice versa.

Remember that Γ-dictators are matched with recipients who received a promise. Again, switching probabilities are reported by row, and the dictators' gender is listed by column.

For instance, the value in the first cell of the table, 0.33, is the average *Roll* rate of female switched Γ-dictators playing the game with a low switching probability (and SOBs high), i.e., $R_F(L, P, S)$. Standard deviations and number of observations are indicated in parentheses.

Let us next look at H3 for men and for women.

Table 2 – *Roll* rates of Γ-dictators

TREATMENT	SWITCHING PROBABILITY	**Women** SWITCH (a)	**Men** SWITCH (b)
(i)	25% (low) [high SOBs]	**0.33** (0.48/45)	**0.50** (0.51/46)
(ii)	75% (high) [low SOBs]	**0.35** (0.48/161)	**0.27** (0.44/161)

The average *Roll* rate of men is significantly higher when the probability of being rematched is low (high SOBs) rather than high (low SOBs): 0.50 vs. 0.27: Z=2.24, p=0.025. Conversely, the average *Roll* rate of switched women is not significantly different in the two matching probability treatments: 0.33 vs. 0.35: Z=-0.26, p=0.398. That is, H3 is confirmed for men but not for women. This supports the idea that GA drives men's motivations to a greater extent than women's ones.

To test H4, we compare the change in *Roll* rates of men for the two different matching probabilities (0.23=0.50-0.27) to the corresponding change in *Roll* rates of women (-0.02=0.33-0.35). H4 is not supported on the 5%-level, although it is close: compare 0.23 vs. -0.02: Z=1.80, p=0.071.

We finally tested for MC (H5-H6). We augmented Table 2 with additional information about the *Roll* rate of not-switched Γ-dictators.

Outcomes are reported in Table 3, which reports the *Roll* rates of switched and non-switched Γ-dictators by gender (columns) for the different probability of switching (rows). Remember that SOBs are likely to be high when the switching probability is low and vice versa. Again, standard deviations and number of observations are indicated in parentheses.

Table 3 – *Roll* rates of Γ-dictators

		Women		**Men**	
TREATMENT	SWITCHING PROBABILITY	SWITCH (a)	NO-SWITCH (b)	SWITCH (c)	NO-SWITCH (d)
(i)	25% (low) [high SOBs]	**0.33** (0.48/45)	**0.52** (0.50/221)	**0.50** (0.51/46)	**0.53** (0.50/205)
(ii)	75% (high) [low SOBs]	**0.35** (0.48/161)	**0.56** (0.50/68)	**0.27** (0.44/161)	**0.60** (0.49/73)

Following Vanberg (2008), we test MC by comparing the average *Roll* rate of non-switched dictators (who made a promise) to that of switched dictators (who made a promise and are rematched with a recipient who received a promise from someone else). We test MC motivation within gender (H5).

We run four tests by combining the two genders and the two switching probabilities (i.e., two levels of SOBs.) H5a (H5b) focuses on women's MC-driven motivation when the switching probability is low (high), while H5c (H5d) focuses on men's MC-driven motivation when the switching probability is low (high).

The outcomes of the tests are listed below.

1. H5a: $R_F(L,P,N) > R_F(L,P,S)$ 0.52 vs. 0.33: *Z*=2.39, *p*=0.016
2. H5b: $R_F(H,P,N) > R_F(H,P,S)$ 0.56 vs. 0.35: *Z*= 2.31, *p*=0.020
3. H5c: $R_M(L,P,N) > R_M(L,P,S)$ 0.53 vs. 0.50: *Z*= 0.08, *p*=0.937
4. H5d: $R_M(H,P,N) > R_M(H,P,S)$ 0.60 vs. 0.27: *Z*= 2.12, *p*=0.034

Our results support MC-driven motivations for women (H5a and H5b), while they support MC for men only when SOBs are low (H5d).

Finally, we test MC sensitivity between genders (H6). We have two cases, one for each SOB level. We look at gender differences by rows. Let us start with the case of low SOBs; if the switching probability is high (row (ii) in Table 3), the data do not provide support for any difference in the MC impact within genders, i.e., (0.60-0.27=0.33) vs. (0.56-0.35=0.21): *Z*=0.18, *p*=0.859. Hence, H6 holds. Now we focus on the case of low SOBs (row (i) in Table 3), we compare the change in *Roll* rates of men (0.52-0.33=0.19) to the corresponding change in *Roll* rates of women (0.53-0.50=0.03). If SOBs are low, H6 is not supported on the 5%-level, although it is close: compare 0.19 vs. 0.03: *Z*=1.73, *p*=0.084.

Our results tend to support that women are sensitive to MC independently of their SOBs, while men are only sensitive to MC when their SOBs are low.

5. Conclusion

We explore guilt aversion (GA) and moral commitment (MC) in a random dictator game with pre-play communication. To the best of our knowledge, except for Kleinknecht's (2019) contribution discussed above, no previous study focused on GA-related gender effects.

Our results support the idea that moral motivations drive women's behavior. By contrast, men are sensitive to SOBs, i.e., the experimental outcomes are consistent with the hypothesis that men are more guilt averse than women. The results also support the idea (with which we opened our paper) that a pregnancy-related biological asymmetry could be the underlying cause: To produce offspring, men need to rely on women's trust, and GA helps make men trustworthy.

Let us critically assess the relevance of that story. A reason to be skeptical is that other situations may favor female GA. For example, Leda Cosmides suggested that a man investing in a woman's offspring (post-birth) has to trust that she has been sexually faithful to him. In this situation, she may benefit more than he does from being guilt averse. Hence, if there are gender differences, those might be domain specific.[9]

Our experiment, involving neither pregnancies nor infidelities, does not decouple any story. We have not explored the relevance of domain-specificity. However, it is possible that evolution in shaping GA does not distinguish various trust scenarios. After all, we have found evidence for GA (among men) in our lab setting, which has no counterpart in our evolutionarily important distant past. Perhaps differences in GA, for a broad class of situations, instead reflect the frequency with which such a trait is evolutionarily advantageous. Based on a rather abstract design, our experiment may then record decision-makers' tendencies more broadly. Perhaps our trust-during-pregnancy scenario was relatively frequent back when evolution shaped our guilt tendencies. Perhaps the infidelity story, favoring female guilt, was relevant less often. [10]

Tooby and Cosmides (2008, p. 117) write that "an emotion is a bet placed under conditions of uncertainty: It is the evolved mind's bet about what internal deployment is likely to lead to the best average long-term set of payoffs, given the structure and statistical contingencies present in the ancestral world." If our abstract lab design carries external validity to many other naturally occurring settings, then, with respect to GA, evolution seems to have placed its bet on the men.

Appendix A

This appendix describes the FOBs and SOBs elicitation procedure. After the communication phase, recipients were asked to guess what their (unknown) dictator would choose to do. They had been told the switching probability (either 75% or 25%). Recipients expressed their guesses by

[9] A few other stories again favor male guilt. See Dufwenberg's (2002) analysis of a "marital investment game" featuring a trusting wife with a guilt-averse husband for reasons other than pregnancy: The "asymmetric treatment of the sexes is consistent with Weitzman's (1986, p. 67) observation that: '[h]usbands and wives typically invest in careers—most particularly in the husband's education and career—and the products of such investments are often a family's major assets,' with Borenstein and Courant's (1989; Footnote 3) observation that a medical student with a supporting spouse typically is a husband with a wife, with evidence concerning divorce cases decided in U.S. courts (Polsby and Zelder 1994, Footnote 4), and with Cohen's (1987) general finding that nuptial gains tend to accrue to men early on in a marriage and to women towards the end."

[10] Alternatively, other emotions may have evolved to make guilt less relevant in the infidelity scenario. For example, male anger, following detected infidelity, may deter infidelity. The evolutionary significance of female guilt proneness would thereby be reduced.

ticking one of the five-point scale in Table A1. This scale is the same as in Vanberg. Beliefs are then re-scaled to 1.00, 0.75, 0.50, 0.25, and 0.00.

Table A1 – Incentives for first-order belief elicitation

The dictator will	choose *Roll*			choose *Don't Roll*	
	Certainly	Probably	Unsure	Probably	Certainly
Please tick your guess	◯	◯	◯	◯	◯
Your earnings					
if the dictator					
chooses *Roll*	0.65 tokens	0.60 tokens	0.50 tokens	0.35 tokens	0.15 tokens
chooses *Don't Roll*	0.15 tokens	0.35 tokens	0.50 tokens	0.60 tokens	0.65 tokens

After dictators were told whether their paired recipient had been switched and read their previous communication, they were asked to guess their guess. Specifically, they had to guess which of the five points of Table A1 had been ticked by their counterpart. Correct guesses earned 0.50 tokens.

Appendix B

The paper focuses on between and within gender clean tests of GA and MC for Γ-dictators. Here, we perform four additional tests akin to Vanberg's (2008) and Kleinknecht (2019) GA test. Recipients who received a promise hold higher FOBs than recipients who did not receive a promise. GA test can then be built by comparing the behavior of dictators whose switched recipients received a promise from someone else with that of dictators whose switched recipients did not receive a promise.

Our data are reported in Table B, where we also consider the *Roll* rates of switched Γ-dictators matched with a recipient who did not receive a promise.

Table B – *Roll* rates of Γ-dictators

TREATMENT	SWITCHING PROBABILITY	MESSAGE	Gender WOMEN (a)	 MEN (c)

(i)	25%	Read (high SOBs)	**0.33** (0.48/45)	**0.50** (0.51/46)
		Do not Read (low SOBs)	**0.18** (0.39/17)	**0.22** (0.42/23)
(ii)	75%	Read (high SOBs)	**0.35** (0.48/161)	**0.27** (0.44/161)
		Do not Read (low SOBs)	**0.30** (0.46/50)	**0.29** (0.46/56)

Table B allows us to compare the behavior of switched Γ-dictators facing a recipient who, everything else equal, received a promise (high SOB) to one who did not (low SOB). However, as discussed in the main text, the two kinds of compared dictators differ alomg two dimensions: 1) SOBs and 2) message read.) Hence, the outcomes of these tests in supporting or not GA can be questionable.

Our findings are as follows.

1. In treatment (i), the data support GA or a different attitude toward the message read between men and women. For the latter, we compare 0.33 (high SOBs) vs. 0.18 (Low SOBs): Z=2.34, p=0.019, while for men, we compare 0.50 (high SOBs) vs. 0.22 (low SOBs): Z=2.30, p=0.022. Differently from Kleinknecht (2019), we find support for men GA.
2. In treatment (ii), no GA support is found. In the case of the women, we compare 0.35 (high SOBs) vs. 0.30 (low SOBs): Z=1.06, p=0.289. In the case of men, we compare 0.27 (high SOBs) vs. 0.29 (low SOBs): Z=0.94, p=0.347.

References

Attanasi, G., C. Rimbaud, and M. C. Villeval (2019), "Embezzlement and guilt aversion," *Journal of Economic Behavior & Organization* 167, 409-429.

Battigalli, P., and M. Dufwenberg (2007), "Guilt in games," *American Economic Review,* 97: 170–176.

Battigalli, P. and M. Dufwenberg (2022), "Belief-dependent motivations and psychological game theory," *Journal of Economic Literature* 60, 833-882.

Bellemare, C., A. Sebald, and S. Suetens (2017), "A note on testing guilt aversion," *Games and Economic Behavior*, 102: 233-239.

Bolton, G. E., and E. Katok (1995), "An experimental test for gender differences in beneficent behavior," *Economics Letters*, 48: 287–292.

Borenstein, S. and P. N. Courant (1989), "How to carve a medical degree: human capital assets in divorce settlements," *American Economic Review,* 79, 992–1009.

Cartwright E. (2019), "A survey of belief-based guilt aversion in trust and dictator games," *Journal of Economic Behavior & Organization,* 167: 430-444.

Charness, G. and M. Dufwenberg (2006), "Promises and partnership," *Econometrica*, 74: 1579–1601.

Childs J. (2012), "Gender differences in lying," *Economics Letters*, 2012, 114: 147–149.

Cohen, L. (1987), "Marriage, divorce, and quasi rents; or, I gave him the best years of my life," *Journal of Legal Studies,* 16. 267–303.

Croson R., and U. Gneezy (2009), "Gender differences in preferences," *Journal of Economic Literature*, 47: 448–474.

Dhami S., M. Wei and A. al-Nowaihi (2019), "Public goods games and psychological utility: Theory and evidence," *Journal of Economic Behavior and Organization*, 167: 361-390.

Di Bartolomeo, G., M. Dufwenberg, and S. Papa (2019a), "The sound of silence: A license to be selfish," *Economics Letters*, 182: 68–70.

Di Bartolomeo, G., M. Dufwenberg, S. Papa, and F. Passarelli (2019b), "Promises, expectations & causation," *Games and Economic Behavior*, 113: 137–146.

Di Bartolomeo, G., M. Dufwenberg, and S. Papa (2022), "Promises and partner-switch," mimeo.

Dreber, A. and M. Johannesson (2008), "Gender differences in deception," *Economics Letters,* 99: 197-199.

Dufwenberg, M. (2002), "Marital investments, time consistency and emotions," *Journal of Economic Behavior & Organization,* 48: 57–49.

Dufwenberg, M. and U. Gneezy (2000), "Measuring beliefs in an experimental lost wallet game," *Games & Economic Behavior* 30, 163-82.

Eckel, C. C., and P. J. Grossman (1998), "Are women less selfish than men? Evidence from dictator experiments," *Economic Journal*, 108: 726–735.

Ellingsen, T., M., Johannesson, S., Tjøtta, and G. Torsvik (2010), "Testing guilt aversion," *Games and Economic Behavior*, 68: 95–107.

Kleinknecht, J. (2019), "A man of his word? An experiment on gender differences in promise keeping," *Journal of Economic Behavior & Organization,* 168: 251–268.

Polsby, D.D. and M. Zelder (1994), "Risk-adjusted valuation of professional degrees in divorce", *Journal of Legal Studies,* 33, 273–285.

Tooby, J. and L. Cosmides (2008), "The evolutionary psychology of the emotions and their relationship to internal regulatory variables," in M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), Handbook of emotions (pp. 114–137). The Guilford Press.

Vanberg, C. (2008), “Why do people keep their promises? An experimental test of two explanations,” *Econometrica*, 76: 1467–1480.

Weitzman, L.J. (1986), “A new look at career assets,” *Ms. Magazine*, Feb., 67–68.

Text of the experiment instructions (Guilt aversion and other motivations: Eve versus Adam)

Instructions

Thank you for participating in this experiment. This experiment aims to study how people make decisions. If you have a question, please raise your hand at any time. However, please do not talk to the other participants; otherwise, you will not be able to earn money.

During the experiment, you will be able to earn tokens based on the decisions you will make (as explained below). In addition, all participants will receive a participation fee of 3 tokens, independent of their choices. Total earnings will be paid to each participant in cash at the end of the experiment. Before starting the experiment, everyone will have to complete a survey.

The experiment consists of 8 independent rounds. In each round, you will interact in a game with a *different* participant randomly chosen by the computer. **You will never interact with the same participant more than once**. At a particular stage of each game, your partner may change. The change can occur with a given probability known from the beginning. This probability (switching probability) can be either high or low. No participant will even know the subject's identity with whom he or she interacted during the experiment. Your payment will be based on your decisions during one of the eight rounds. That round will be randomly selected at the end of the experiment (each round has the same probability of being selected.) Each round consists of 5 steps, described below.

Step 1: Communication

At the beginning of each round, you will be randomly paired with a partner and have the opportunity to communicate with him/her by exchanging messages through the PC. The computer will randomly choose the first participant in the pair to send a message; then, participants will alternate. Each participant will be able to send up to two messages.

Important. Regarding the content of the messages, you are not allowed to reveal your identity to your paired participant. (You should not reveal your name or any other identification feature, such as sex, hair color, the location of your computer, etc.) Other than these exceptions, you can send whatever message you want. Please continue to be silent during the experiment. Participants who violate these rules (at the discretion of the controllers) will be excluded from the experiment.

Step 2: Random assignment of the role

After communication, you will randomly be assigned to either role A or role B (based on your role, you will be asked to perform different tasks). If you are assigned role A, your partner during communication will take role B and vice versa. You will play the whole round in your assigned role; the computer will reassign roles in the subsequent round. Note that the probability of being assigned either role A or role B is always the same. It does not depend on who started the first communication in step 1. It means that, at the time of communication, no participant has received any information about his or her role, either A or B.

Step 3: Switching probability

Hereafter, in any round, the computer will show everyone the probability of being re-matched with a different partner than the one you previously communicated with. **The probability that you will interact with a different subject can either be low (25%) or high (75%).**

Step 4: Possible partner switch

If a partner switch occurs, in this step, participant A will know whether his/her partner was switched or not. That switch is such that any new switched pair will be composed of one participant A and one B. Note that only participant A will know whether or not her partner has been switched, while participant B will never get to know it.

Step 5: A's choice

During this decision phase, Participant B chooses whether to roll or not a six-sided die. If B chooses not to roll it. B receives 14 tokens, and A receives 0 tokens. If B chooses to roll the dice, B receives 10 tokens and the PC will randomly draw a number between 1 and 6 that determines the payoff of A. If number 1 is drawn, A receives 0 tokens. All numbers from 1 to 6 are equally likely, as in a standard dice. If numbers from 2 to 6 are drawn, A receives 12 tokens. All earnings are summarized in the table below (you must add the participation fee of 3 tokens, which is independent of your choices).

	Player A earns	**Player B earns**
If A chooses not to roll the dice	**14 Tokens**	**0 Tokens**
If A chooses to roll the dice and the outcome is 1	**10 Tokens**	**0 Tokens**
If A chooses to roll dice and the outcome is 2, 3, 4, 5, 6	**10 Tokens**	**12 Tokens**

Information at the end of each round.

Participant A will be informed about the earnings of both participants at the end of each round. If A chooses to roll the dice, A will observe the outcome. Participant B will only be informed about his/her earnings. Thus, if you are selected to play role B and earn 0 tokens, you will never know what choice A made. The reason is that B earns 0 tokens in two cases: a) if A has chosen not to roll the dice; b) if A chooses to roll the dice, and the outcome is 1. Remember that **B will never know whether he/she has been switched after the communication phase. B only knows whether the probability of being switched is high (75%) or low (25%).**

Additional earnings:

At some point in the experiment, you will have the opportunity to earn additional tokens if you answer some questions correctly. For instance: "Guess your payoffs" or "Guess what the other subject expects about his/her payoffs."

You will always be paid when your guesses are correct, except in the round that the computer will draw for payment. In that round, you will not be paid for your guesses. You will only be paid based on the dice outcome.

EXCHANGE RATE: 1 TOKEN = 0.50 EURO